JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **136**, 466–478 (1988)

About When to Use the Searchlight

GEERT JAN OLSDER

Department of Mathematics and Informatics,
Delft University of Technology,
P.O. Box 356, 2600 AJ Delft, The Netherlands

AND

GEORGE P. PAPAVASSILOPOULOS*

Department of Electrical Engineering Systems,
University of Southern California,
Los Angeles, California 90089–0781

Submitted by S. M. Meerkov

Received November 3, 1986

Two players, P and E, not knowing each others' positions, move in a domain. Player P has a searchlight which he can flash at will and which then illuminates a certain area around P. The game ends when E is caught within this area, provided it is illuminated. If E is not in the illuminated area, then P has disclosed his position to E since E can observe the searchlight, if it is switched on, from everywhere. P wants to maximize and E wants to minimize the capture chance over a given time horizon. This paper provides a dynamic programming formulation of this game, which in its turn yields optimal strategies for the players, i.e., how to move and for P, in addition, at which time instants to flash. The game is considered on a finite state space and in discrete time. © 1988 Academic Press, Inc

1. INTRODUCTION AND PROBLEM STATEMENT

The game to be discussed belongs to the class of two person zero-sum dynamic games of the pursuit–evasion type. Two players, called P and E, move in a certain domain and are unaware of each others' positions unless P flashes a searchlight which illuminates an area of known shape around this player and which discloses P's position to E. Termination of the game occurs if P flashes his searchlight and E finds himself trapped within the

* The work of this author was supported by ZWO, the Netherlands Organization for the Advancement of Pure Research, Number B62–239, while he spent a sabbatical at Delft University of Technology,

0022-247X/88 $3.00

area illuminated. Flashing the searchlight has two aspects; it is the only way for P to terminate the game and it discloses P's position to E. This latter aspect may be used by E in order to minimize the capture chance later on during the game (if he is not captured immediately) since P may flash again.

It is easy to understand that both players will use probabilistic strategies for their dynamic behaviour in the domain since a pure strategy by one player would certainly lead to a loss for him (if initially the players know each others' positions and if P would have an optimal pure strategy, E can calculate this strategy also and therefore avoid P; if E would use an optimal pure strategy, then P can calculate this strategy also and certainly capture E if the time horizon is sufficiently long). Conceptually it is very difficult to define mixed or behavioural strategies for games which proceed continuously in time. Therefore the problem has been investigated for a finite state space, i.e., the players move in a network with a finite number of nodes, and for the discrete time case. It is possible that both players occupy the same node in the network at the same time without realizing this. Only if P flashes at such an instant do the players become aware of this coincidence.

In this paper we confine ourselves to a specific network; the nodes are positioned on the circumference of a circle and each node has two adjacent nodes; one to the right and one to the left. During each time step each player can stay where he is, move one node to the left or one to the right. Conceptually, generalizations to other networks are not difficult; the notation, however, would become more cumbersome. At each time step P has the option whether to flash or not. The illuminated area consists of three nodes; the node at which P is situated, the one to the left and the one to the right. Given a finite time horizon, P wants to maximize and E wants to minimize the capture probability during this time interval.

A practical motivation for this game might be the following. E is a smuggler who wants to steer his boat to the shore and P is a police boat. For obvious reasons E wants to avoid P and therefore he performs his practices only during the night and that is exactly why the police boat will use a searchlight.

A variation of this game has been described in [1]. In [1] both players had a searchlight and both tried to capture the other player before they were captured themselves. A serious limiation of [1] was that both players could use their searchlights only once. Finding the optimal strategies boiled down to the solution of a hierarchical linear programming problem, i.e., the outcome of one linear programming problem was used as an entry to another linear programming problem. It was shown that the same solution method could be used for a different problem where P had two flashes at his disposal and E none. The current paper can thus be viewed as a generalization of this latter problem. In order to find the solution one must now solve a series of hierarchical nonlinear programming problems which are a consequence of the dynamic programming approach.

2. The Basic Game: P Has One Flash at His Disposal

The basic game which is a building block for the general game is formulated as follows. Initially P and E know each others' positions on the circle; P is at node p and E is at node e; p, $e \in \{1, ..., n\}$, where n is the number of nodes on the circumference of the circle. The numbering of the nodes is in chronological order; node n is again adjacent to node 1. Player P can use his searchlight only once during the time interval $t=1$, $t=2, ..., t=T$; the final time T is fixed and known to both players. Once the time proceeds the players do not get to know about their opponent's new positions (unless P flashes). It has been shown in [1] that the optimal strategies of the basic game can be found from the matrix game

$$\min_{y_e \in S_{3^T}} \max_{x_p \in S_{d_T}} y_e' A_{p,e,T} x_p. \tag{1}$$

The meaning of the various symbols introduced is as follows. The symbol $'$ denotes transpose. At each instant of time E can choose from three options (move to one of the two adjacent nodes or stay where he is) and therefore his number of pure strategies is 3^T. Each component of the vector y_e denotes the probability with respect to which such a pure strategy will be chosen. The set S_{3^T} is the simplex to which the 3^T-vector y_e belongs; all components are nonnegative and add up to one. The components of the vector x_p indicate P's pure strategies, determined by when and where to flash. Vector x_p has d_T components. For $T \leqslant [n/2]$ it has been shown in [1] that $d_T = T^2 + 2T$. The set S_{d_T} denotes the simplex from which x_p must be chosen. The elements of the matrix $A_{p,e,T}$ are either one (if for the particular strategies capture takes place) or zero (if no capture takes place). The first index of A refers to P's initial position, the second index to E's initial position, and the third index refers to the number of time steps during which the flash will take place.

It has been shown in [1] that, if the saddle-point value of (1) is indicated by $J_{p,e,T}$,

$$J_{p,e,1} \leqslant J_{p,e,2} \leqslant \cdots \leqslant J_{p,e,[n/2]} = J_{p,e,[n/2]+1} = J_{p,e,[n/2]+2} = \cdots,$$
$$p, e \in \{1, ..., n\}.$$

Therefore one can restrict oneself to a time horizon of $[n/2]$ steps in order to solve the game. The idea behind the proof is that in $[n/2]$ steps each player can reach any point of the circle, with an arbitrary underlying probability distribution.

For later reference the matrix $A^0_{p,e,T}$ is introduced which has the same size as matrix $A_{p,e,T}$ and is constructed in the same way except for the fact that for $A^0_{p,e,T}$ capture is defined to take place only if the positions of P and E coincide during the flash (i.e., the illuminated area is P's position only and does not include the adjacent nodes).

Consider the following generalization of the basic game in that at $t=0$ E still knows where P is (at node p), but P only knows the probability distribution of E's position. The probability for E to be (at $t=0$) at node e is v_e. If $v=(v_1, ..., v_n)$, then $v \in S_n$. The vector v is assumed to be known to both P and E. Of course E knows his own position at $t=0$. This game can again be formulated as a matrix game

$$\min_{y_e \in S_{3^T},\, e=1, ..., n} \; \max_{x_p \in S_{d_T}} \left[\sum_{i=1}^{n} v_i y_i' A_{p,i,T} x_p \right]. \tag{2}$$

THEOREM 1. *The minmax problem* (2) *can be solved by means of the following linear programming problem*:

$$(\tilde{y}_1', ..., \tilde{y}_n') \begin{bmatrix} v_1 A_{p,1,T} \\ \hline \vdots \\ \hline v_n A_{p,n,T} \end{bmatrix} \leqslant (1, 1, ..., 1);$$

$$\sum_{j=1}^{3^T} (\tilde{y}_1)_j = \sum_{j=1}^{3^T} (\tilde{y}_i)_j, \qquad i=2, ..., n; \tag{3}$$

$$(\tilde{y}_i)_j \geqslant 0, \qquad i=1, ..., n; j=1, 2, ..., 3^T;$$

$$\max \sum_{j=1}^{3^T} (\tilde{y}_1)_j.$$

Proof. Obviously E wants to minimize $m \in \mathbb{R}$, subject to

$$(y_1', ..., y_n') \begin{bmatrix} v_1 A_{p,1,T} \\ \hline \\ \hline v_n A_{p,n,T} \end{bmatrix} \leqslant m(1, 1, ..., 1), \qquad y_e \in S_{3^T}, e=1, ..., n. \tag{4}$$

The index of the largest component of the vector on the left-hand side of (4) determines the optimal x_p-strategy for P; P will use the pure strategy which corresponds to this index and the outcome of the game equals the value of this largest component. Player E wants to minimize this largest component. By defining $\tilde{y}_e = y_e/m$, the statement of the theorem is readily obtained. A tacit assumption made in this proof is that the minimizing m is positive. This is not a serious restriction, however; see for instance [2] for the modification required in this case. ∎

The maximal value of the criterion in the linear programming problem above equals the inverse of the saddle-point value of (2) which is denoted by $f_{p,1,T}$;

$$f_{p,1,t}(v) = \min_{y_e, e=1,\ldots,n} \max_{x_p} \left[\sum_e v_e y_e' A_{p,e,t} x_p \right], \tag{5}$$

with $t = T$. The first index of f refers to P's initial position, the second index to the number of flashes available to P, and the third index to the time horizon. The expected outcome of the game for P is $f_{p,1,T}(v)$. The outcome for E is $y_e^{*\prime} A_{p,e,T} x_p^*$, where e denotes E's initial position and the starred vectors are the minmax arguments of (5).

THEOREM 2. *The function $f_{p,1,T}(v)$, $v \in S_n$, as a function of v, is convex.*

Proof. If one solves for the minimal y_i-values in (2), expressing them as functions of x_p, the result can be written as

$$f_{p,1,T}(v) = \max_{x_p \in S_{d_i}} \sum_{i=1}^{n} v_i y_i'(x_p) A_{p,i,T} x_p,$$

which is symbolically written as

$$f_{p,1,T}(v) = \max_x \sum_{i=1}^{n} v_i \bar{f}_{p,i}(x).$$

For each x the function $\sum v_i \bar{f}_{p,i}$ is linear with respect to v. The maximum of a set of linear functions, indexed by x, is obviously convex. ∎

EXAMPLE. $f(v) = \min_{y_1, y_2} \max_x [v_1 y_1' A_1 x + v_2 y_2' A_2 x]$, where

$$x = \begin{pmatrix} x_1 \\ 1 - x_1 \end{pmatrix}, y_1 = \begin{pmatrix} y_{11} \\ 1 - y_{11} \end{pmatrix}, y_2 = \begin{pmatrix} y_{21} \\ 1 - y_{21} \end{pmatrix}, v_2 = 1 - v_1;$$

$$0 \leqslant x_1, y_{11}, y_{21}, v_1 \leqslant 1; A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix}.$$

Solving for the minimizing y_{11}, y_{21}-values as functions of x and substituting those into $f(v)$, one obtains

$$y_{11}(x)=\begin{cases}1, & 0\leqslant x<\frac{1}{2};\\ 0, & \frac{1}{2}<x\leqslant 1;\end{cases}\quad y_{21}(x)=\begin{cases}0, & 0\leqslant x<\frac{4}{5};\\ 1, & \frac{4}{5}<x\leqslant 1;\end{cases}$$

$$\min_{y_1, y_2}\,[v_1 y_1' A_1 x+(1-v_1)\,y_2' A_2 x]=\begin{cases}x_1, & 0\leqslant x_1\leqslant\frac{1}{2};\\ v_1+x_1-2v_1x_1, & \frac{1}{2}\leqslant x_1\leqslant\frac{4}{5};\\ (1-x_1)(4-3v_1), & \frac{4}{5}\leqslant x_1\leqslant 1.\end{cases}$$

For each x_1 the functions on the right-hand side are linear in v. For $0\leqslant v_1\leqslant\frac{1}{2}$ the maximum with respect to x_1 is achieved for $x_1=\frac{4}{5}$ which leads to $f(v)=\frac{4}{5}-\frac{3}{5}v_1$. For $\frac{1}{2}\leqslant v_1\leqslant 1$ the maximum is achieved for $x=\frac{1}{2}$ which leads to $f(v)=\frac{1}{2}$. The function $f(v)$ is piecewise linear, continuous, and convex.

If initially the players are placed, independently of each other, on the circumference of the circle according to the uniform distribution, then $v_i=1/n$, $i=1, ..., n$. There is no unique time instant at which P should switch on his searchlight. In fact, he can flash at any time instant $t\in\{1, ..., T\}$, since at each t the players are still distributed according to a uniform distribution. The capture chance therefore equals $3/n$, where 3 is the size of the illuminated area and n the size of the state space. Hence

$$f_{p,1,T}(1/n, ..., 1/n)=1/n\min_{y_e,\,e=1,\dots,n}\,\max_{x_p}\left[\sum_e y_e' A_{p,e,T}x_p\right]=3/n.$$

3. P Must Flash Two Times

As in the previous section, two different games will be distinguished. First the case is considered where initially the players are situated on the circle according to a uniform distribution, independently of each other. In the second case this is not true; there initially P is at position p and E is distributed according to the probability vector $v\in S_n$. Both p and v are known to the players, but only E knows his own position e. In both cases it is assumed that P must flash two times, and not that P may flash two times maximally. The latter case will be dealt with in Section 5.

Case 1. It is no restriction to assume that P flashes for the first time at $t=1$, at his position indicated by p. The nodes $p-1$, p, and $p+1$ are illuminated and the capture chance is $3/n$. If E is not caught at the first flash, then P assumes a uniform distribution for E on the remaining nodes

$1, ..., p-2, p+2, ..., n$, each with probability $1/(n-3)$. The remaining game to be solved by the players is

$$\min_{\substack{y_1, \dots, y_{p-2} \\ y_{p+2}, \dots, y_n}} \max_{x_p} \left[\sum_{\substack{e=1, \dots, p-2 \\ e=p+2, \dots, n}} y'_e A_{p,e,T-1} x_p \right],$$

the outcome of which can be written as

$$f_{p,1,T-1}\left(\frac{1}{n-3}, ..., \frac{1}{n-3}, 0, 0, 0, \frac{1}{n-3}, ..., \frac{1}{n-3}\right),$$

where the zeroes are the $(p-1)$st, pth, and $(p+1)$st arguments. The summation $e=1, ..., p-2, p+2, ..., n$ used in the above formula only makes sense, strictly speaking, for $p=3, ..., n-2$. What is meant is

$$\sum_{\substack{e=1, \dots, n \\ e \neq p-1, p, +1 \text{ modulo } n}},$$

which makes sense for any $p \in (1, ..., n)$. This convention is also used in the remainder of the paper. The total capture probability is

$$\frac{3}{n} + \left(1 - \frac{3}{n}\right) f_{p,1,T-1}\left(\frac{1}{n-3}, ..., \frac{1}{n-3}, 0, 0, 0, \frac{1}{n-3}, ..., \frac{1}{n-3}\right),$$

which is independent of p, due to the symmetry of the nodes on the circle and due to the initial uniform distributions.

Case 2. Since P must flash twice, the first flash must take place at $t \in \{1, ..., T-1\}$. This flash can take place at $t=T-1$ at the latest. Therefore we must, for E, consider 3^{T-1} different pure strategies and this for each of the initial positions $\{1, ..., n\}$. It is assumed that E may play a mixed strategy and the probabilities for the $n * 3^{T-1}$ different strategies are indicated by

$$y_{e;m_1, \dots, m_{T-1}}; \qquad y_e = (y_{e;L,\dots,L}, y_{e;L,\dots,L,M}, y_{e;L,\dots,L,R}, \dots, y_{e;R,\dots,R}) \in S_{3^{T-1}}. \tag{6}$$

The index m_t describes E's move at time t (*L*eft, *M*iddle, or *R*ight node). Index e is E's position at $t=0$.

P's probabilities for possible strategies up to and including the first flash are enumerated according to $x_{p,q,j}$, where the first index refers to P's initial node (at $t=0$), the second index to the node at which P is situated when he switches on the searchlight, and the third index $j \in \{1, ..., T-1\}$ denotes the time at which this occurs. We write

$$x_p = (x_{p,1,1}, x_{p,2,1}, ..., x_{p,n,1}, x_{p,1,2}, ..., x_{p,n,T-1}) \in S_{d_{T-1}}. \tag{7}$$

The quantity $w_{p,e,l}$ is introduced to denote the probability that E will be at node e at time l conditioned on P's knowledge. Therefore $w_{p,e,l}$ equals the sum of a subset of elements out of $\{y_{i;m_1,\ldots,m_{T-1}}, i=1,\ldots,n;\ m_j=L,\ M$ or $R\}$, divided by n. Obviously

$$\sum_{e=1,\ldots,n} w_{p,e,l}=1, \qquad l=1,\ldots,T-1; \tag{8}$$

$$w_{p,e,l}=\sum_{i=1}^{n} v_i y_i' A^0_{p,i,T-1} u_{p,e,l}, \tag{9}$$

where v_i corresponds to E's initial distribution $v=(v_1,\ldots,v_n)$, $A^0_{p,i,t}$ was introduced in Section 2, and where $u_{p,i,l}$ denotes the x_p-vector with all elements being zero except for the element indexed by p, i, l (see (7)), which equals one.

The capture chance when P flashes for the first time, which could be any time instant between 1 and $T-1$, is

$$\sum_{i=1}^{n} v_i y_i' A_{p,i,T-1} x_p=\sum_{l=1}^{T-1}\sum_{q=1}^{n} x_{p,q,l}(w_{p,q-1,l}+w_{p,q,l}+w_{p,q+1,l}). \tag{10}$$

If E is not caught during the first flash, the probability that E will be caught during the second flash is, provided both P and E play optimally from the time instant of the first flash (which took place at time l and node q) onwards, is

$$f_{q,1,T-l}\left(\frac{w_{p,1,l}}{\sum_{i=1,\ldots,q-2,i=q+2,\ldots,n} w_{p,i,l}},\ldots,\frac{w_{p,q-2,l}}{\text{same denominator}},0,0,0,\frac{w_{p,q+2,l}}{\text{same denominator}},\ldots,\frac{w_{p,n,l}}{\text{same denominator}}\right). \tag{11}$$

The total capture probability, with respect to both flashes, is

$$\begin{aligned}&\sum_{l=1}^{T-1}\sum_{q=1}^{n} x_{p,q,l}(w_{p,q-1,l}+w_{p,q,l}+w_{p,q+1,l})\\&\quad+\left(1-\sum_{l=1}^{T-1}\sum_{q=1}^{n} x_{p,q,l}(w_{p,q-1,l}+w_{p,q,l}+w_{p,q+1,l})\right)\\&\quad\times\left(\sum_{l=1}^{T-1}\sum_{q=1}^{n} x_{p,q,l} f_{q,1,T-l}\left(\frac{w_{p,1,l}}{\sum_{i=1,\ldots,q-2,i=q+2,\ldots,n} w_{p,i,l}},\right.\right.\\&\quad\ldots,\frac{w_{p,q-2,l}}{\text{same denominator}},0,0,0,\\&\quad\left.\left.,\frac{w_{p,q+2,l}}{\text{same denominator}},\ldots,\frac{w_{p,n,l}}{\text{same denominator}}\right)\right).\end{aligned} \tag{12}$$

This expression must be maximized with respect to x_p and minimized with respect to $y_1, \ldots, y_n$ and the resulting outcome is denoted by

$$f_{p,2,T}(v) = \min_{y_1, \ldots, y_n} \max_{x_p} (12). \tag{13}$$

Expression (13) is the first step of the dynamic programming approach, working backwards in time from one to two flashes. The value function is indicated by the function f, the state for the dynamic programming approach is (p, v). In order to solve (13), the total time horizon must be known, however, as well as the number of flashes. The latter quantities can be viewed as parameters of the system under consideration. Equation (12) can be simplified slightly to

$$\begin{aligned}
&\sum_{l=1}^{T-1} \sum_{q=1}^{n} x_{p,q,l}(w_{p,q-1,l} + w_{p,q,l} + w_{p,q+1,l}) \\
&\quad + \left(\sum_{l=1}^{T-1} \sum_{q=1}^{n} x_{p,q,l} \left(\sum_{i=1,\ldots,q-2,\, i=q+2,\ldots,n} w_{p,i,l} \right) \right) \\
&\quad \times \left(\sum_{l=1}^{T-1} \sum_{q=1}^{n} \frac{w_{p,q,l}}{\sum_{i=1,\ldots,q-2,\, i=q+2,\ldots,n} w_{p,i,l}} \right. \\
&\quad \left. \times f_{q,1,T-l}(w_{p,1,l}, \ldots, w_{p,q-2,l}, 0, 0, 0, w_{p,q+2,l}, \ldots, w_{p,n,l}) \right).
\end{aligned} \tag{14}$$

An important question is whether the saddle point as expressed by (13) does exist. We would like to show that (12), or (14), is concave in x_p and convex in $(y_1, \ldots, y_n)$. Though this is very likely to be true—the results of the following two lemma's point already in this direction—a general proof is not currently available.

LEMMA 1. *Expression* (12) *is convex in* $(y_1, \ldots, y_n)$ *if*

(i) x_p *is a pure strategy*;

(ii) $x_{p,e,t} = 1/(n * (T-1))$, $e = 1, \ldots, n$; $t = 1, \ldots, T-1$.

Proof. (i) The first term in (12) is linear in $w_{p,e,t}$ and therefore linear in $(y_1, \ldots, y_n)$. Hence only the convexity of the second term still needs to be shown. If all $x_{p,q,r}$-terms are zero except one, say $x_{r,s,m}$, then this second term becomes

$$\begin{aligned}
&(1 - w_{r,s-1,m} - w_{r,s,m} - w_{r,s+1,m}) \\
&\quad \times f_{s,1,T-m}\left(\frac{w_{r,1,m}}{\sum_{i=1,\ldots,s-2,\, i=s+2,\ldots,n} w_{r,i,m}}, \ldots \right) \\
&\quad = f_{s,1,T-m}(w_{r,1,m}, \ldots, w_{r,s-2,m}, 0, 0, 0, w_{r,s+2,m}, \ldots, w_{r,n,m})
\end{aligned}$$

which is convex in its arguments (see Theorem 2). Since the arguments are linear in the components of $(y_1, ..., y_n)$, $f_{s,1,T-m}$ is also convex in $(y_1, ..., y_n)$ which concludes part (i) of the lemma.

(ii) As in part (i), the first term in (12) is linear in $w_{p,e,l}$ and only the convexity of the second term still needs to be shown. With $x_{p,e,l} = 1/(n*(T-1))$ this second term becomes

$$\left(1-\frac{3}{n}\right) \times \frac{1}{n(T-1)} \left(\sum_{l=1}^{T-1} \sum_{q=1}^{n} f_{q,1,T-l}(\cdots) \right).$$

Because of the symmetry of the nodes along the circle circumference, and because of the definition of $f_{q,1,t}$ in (5),

$$\sum_{q=1}^{n} f_{q,1,T-l}(\cdots) = \sum_{i=1}^{n} \sum_{q=1}^{n} c_{iq} \frac{w_{p,i,l}}{\sum_{j=1,\ldots,q-2,j=q+2,\ldots,n} w_{p,j,l}},$$

where the c_{iq} are nonnegative coefficients. Since $w_{p,i,l}$ is linear in the elements of $(y_1, ..., y_n)$, also with nonnegative coefficients, it now easily follows that

$$\sum_{q=1}^{n} f_{q,1,T-l}(\cdots)$$

is convex in the coefficients of $(y_1, ..., y_n)$, which proves the second part of the lemma. ∎

LEMM 2. *Expression* (12) *is linear, and therefore concave, in* x_p *if* $w_{p,e,l} = (1/n)$, $e = 1, ..., n$.

Proof. The first term in (12) is linear in $x_{p,q,l}$. The second term can be written as

$$\begin{aligned} &\left(\left(1-\frac{3}{n}\right) \sum_{l=1}^{T-1} \sum_{q=1}^{n} x_{p,q,l}\right) \\ &\qquad \times \left(\sum_{l=1}^{T-1} \sum_{q=1}^{n} x_{p,q,l} f_{q,1,T-l} \left(\frac{1}{n-3}, \right.\right. \\ &\qquad \left.\left. \ldots, \frac{1}{n-3}, 0, 0, 0, \frac{1}{n-3}, \ldots, \frac{1}{n-3} \right)\right) \\ &= \left(1-\frac{3}{n}\right) \left(\sum_{l=1}^{T-1} \sum_{q=1}^{n} x_{p,q,l} f_{q,1,T-l} \right), \end{aligned}$$

which proves the lemma. ∎

It must be remarked that it is may be not possible to choose $(y_1, ..., y_n)$ in such a way that $w_{p,e,l} = 1/n$ for all indices p, e, l. However, whenever it is possible, Lemma 2 shows the concavity. Also,

$$f_{q,1,T-l}\left(\frac{1}{n-3}, ..., \frac{1}{n-3}, 0, 0, 0, \frac{1}{n-3}, ..., \frac{1}{n-3}\right)$$

with the zeroes as the $(q-1)$st, qth, and $(q+1)$st elements, is independent of q, but this fact is not needed in the proof of the lemma.

It may very well be possible that the moment of the first flash is not a random variable, but instead a deterministic one. If this could be shown many formulas above could be simplified since a mixing with respect to time (the running index l) would not be necessary anymore.

4. The General Case: P Must Flash K Times

Initially P is at node p and E is distributed according to a probability vector $v \in S_n$. Both p and v are known to the players, as well as the number of flashes K and the time horizon $T \geqslant K$. Player P also knows his own initial position e.

It follows directly from Section 3 that $f_{p,k,t}(v)$, the saddle-point value if there are t steps and k flashes to go and the initial situations of P and E are indicated by p and v, respectively, is recursively determined by increasing k,

$$\begin{aligned} f_{p,k,t}(v) = \min_{\substack{y_1, ..., y_n \\ y_e \in S_3^{t-k+1}}} \max_{x_p \in S_{d_{t-k+1}}} \Bigg[& \sum_{i=1}^{n} v_i y_i' A_{p,i,t-k+1} x_p \\ & + \left(1 - \sum_{i=1}^{n} v_i y_i' A_{p,i,t-k+1} x_p\right) \\ & \times \Bigg(\sum_{l=1}^{t-k+1} \sum_{q=1}^{n} x_{p,q,l} f_{q,k-1,t-l} \\ & \left(\frac{w_{p,1,l}}{\sum_{i=1,...,q-2,i=q+2,...,n} w_{p,i,l}}, ..., \frac{w_{p,q-2,l}}{\text{same denominator}}, 0, 0, 0, \right. \\ & \left. , \frac{w_{p,q+2,l}}{\text{same denominator}}, ..., \frac{w_{p,n,l}}{\text{same denominator}}\right)\Bigg)\Bigg], \end{aligned} \tag{15}$$

$t = k, ..., T$; $p = 1, ..., n$,
where

$$w_{p,r,l} = \sum_{i=1}^{n} v_i y_i' A^0_{p,i,t-l-1} u_{p,r,l}.$$

It is conjectured that for given n, T, p, and v, $f_{p,K,T}(v)$, as a function of K, has only one maximum. If K is close to zero or K is close to T, $f_{p,K,T}(v)$ will be small. For $K \sim T/n$ or slightly greater, $f_{p,K,T}(v)$, as a function of K, will be maximal.

It is also conjectured that the expression between square brackets in (15) is convex in $(y_1, ..., y_n)$ and concave in x_p. In addition it is conjectured that $f_{p,k,t}(v)$ is convex in v. If one would assume the convexity of $f_{p,k-1,t}(v)$, $p = 1, ..., n$, $t = k-1, ..., T$, then the reasoning of Lemmas 1 and 2 applies in order to show that the above-mentioned expression in square brackets is indeed convex/concave in some specific cases.

5. P Has K Flashes at His Disposal

The problem to be considered is the same as in Section 4, with one exception and that is that P does not have to flash K times; K only denotes the maximum number of flashes that K has at his disposal. If the saddle-point value is now indicated by $g_{p,K,T}(v)$, provided it exists, it is recursively determined by

$$\begin{aligned} g_{p,k,t}(v) = \min_{\substack{y_i \in S_{3T} \\ i=1,...,n}} \max_{x_p \in S_{d_t}} \Bigg[& \sum_{i=1}^{n} v_i y_i' A_{p,i,t} x_p \\ & + \left(1 - \sum_{i=1}^{n} v_i y_i' A_{p,i,t} x_p\right) \\ & \times \Bigg(\sum_{l=1}^{t} \sum_{q=1}^{n} x_{p,q,l}\, g_{q,\min(k-1,t-l),t-l} \\ & \left(\frac{w_{p,1,l}}{\sum_{i=1,...,q-2,i=q+2,...,n} w_{p,i,l}}, ..., \frac{w_{p,q-2,l}}{\text{same denominator}}, 0, 0, 0,\right. \\ & \left.\left. , \frac{w_{p,q+2,l}}{\text{same denominator}}, ..., \frac{w_{p,n,l}}{\text{same denominator}}\right)\Bigg)\Bigg], \end{aligned} \tag{16}$$

$t = k, ..., T; p = 1, ..., n$,
where now

$$w_{p,r,l} = \sum_{i=1}^{n} v_i y_i' A^0_{p,i,t} u_{p,r,l}.$$

The essential difference of (16) as compared to (15) is that the next flash may take place at the last time instant $t = T$ and therefore the size of the strategy spaces must be adjusted accordingly. As long as E is not caught, $k < K$ and $t < T$, there will always be another flash since that yields another nonnegative probability of capture.

6. Conclusions

We have discussed a discrete-time game with dynamic information in a finite state-space. For reasons of simplicity we have considered the state-space consisting of n elements which are positioned around the circumference of a circle. Extensions of the state-space to more general networks of nodes are possible and the techniques developed in this paper seem to be extendable to those more general (finite) state-spaces, though the notation will become more complicated. Two players move in the state-space, in discrete time, and one of the players, P, can terminate the game by catching the other player, E, in his searchlight. Switching on the searchlight, however, also discloses P's position to E; who, if not caught at the same moment, can use this information during the remainder of the game.

Dynamic programming equations have been derived for this game. The "state" for these equations turns out to be P's initial position, the distribution of E with respect to the nodes as viewed by P, the number of flashes still available, and the number of time steps to go. One step in the dynamic programming formulation corresponds to going from one time instant at which the light was switched on to the next time instant at which the light was switched on. These time instants are in principle stochastic events, determined by the dynamic programming equations. The total time duration of the game is a parameter in the dynamic programming equations. This is not standard since one does not simply work backward in time, but one has to know the initial time. Existence questions as to whether a solution to the dynamic programming equations exists have been discussed. The proof for a general existence seems to be difficult and only some limited results were shown to hold true.

A possible extension of the game discussed is the following one. Both players have a flashlight at their disposal which they may switch on any moment they like. Each player tries to capture the other one before he is caught himself. This would be a generalization of the game considered in [1] where each player could switch on his searchlight only once.

References

1. G. J. Olsder and G. P. Papavassilopoulos, A Markov chain game with dynamic information, *J. Optim. Theory Appl.* **59** (1988), 467–486.
2. T. Basar and G. J. Olsder, "Dynamic Noncooperative Game Theory," Academic Press, London, 1982.